

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: BURRELL, MARILEE

HILL, DAVID E.

KINZLER, KENNETH W.

VOGELSTEIN, BERT

(ii) TITLE OF INVENTION: AMPLIFICATION OF HUMAN MDM2 GENE IN
HUMAN TUMORS

(iii) NUMBER OF SEQUENCES: 5

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: BANNER, BIRCH, MCKIE AND BECKETT

(B) STREET: 1001 G STREET, N.W.

(C) CITY: WASHINGTON

(D) STATE: D.C.

(E) COUNTRY: USA

(F) ZIP: 20001

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk

(B) COMPUTER: IBM PC compatible

(C) OPERATING SYSTEM: PC-DOS/MS-DOS

(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

00965724-100101

(vi) CURRENT APPLICATION DATA:

- (A) APPLICATION NUMBER:
- (B) FILING DATE: 07-APR-1993
- (C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: KAGAN, SARAH A.
- (B) REGISTRATION NUMBER: 32,141
- (C) REFERENCE/DOCKET NUMBER: 01107.42798

(ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: 202-508-9100
- (B) TELEFAX: 202-508-9299
- (C) TELEX: 197430 BBMB UT

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 64 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

096674 100T 42293560

(v) FRAGMENT TYPE: N-terminal

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Homo sapiens

(viii) POSITION IN GENOME:

(A) CHROMOSOME/SEGMENT: 17q

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Met Glu Glu Pro Gln Ser Asp Pro Ser Val Glu Pro Pro Leu Ser Gln
1 5 10 15

Glu Thr Phe Ser Asp Leu Trp Lys Leu Leu Pro Glu Asn Asn Val Leu
20 25 30

Ser Pro Leu Pro Ser Gln Ala Met Asp Asp Leu Met Leu Ser Pro Asp
35 40 45

Asp Ile Glu Gln Trp Phe Thr Glu Asp Pro Gly Pro Asp Glu Ala Pro
50 55 60

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2372 base pairs

(B) TYPE: nucleic acid

0906724 100101

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Homo sapiens

(H) CELL LINE: CaCo-2

(viii) POSITION IN GENOME:

(B) MAP POSITION: 12q12-14

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 312..1784

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GCACCGCGCG AGCTTGGCTG CTTCTGGGGC CTGTGTGGCC CTGTGTGTCG GAAAGATGGA 60

GCAAGAAGCC GAGCCCGAGG GGCGGCCGCG ACCCCTCTGA CCGAGATCCT GCTGCTTTCG 120

CAGCCAGGAG CACCGTCCCT CCCCGGATTA GTGCGTACGA GCGCCCAGTG CCCTGGCCCCG 180

09565724 100101

GAGAGTGGAA TGATCCCCGA GGCCAGGGC GTCGTGCTTC CGCAGTAGTC AGTCCCCGTG 240

AAGGAAACTG GGGAGTCTTG AGGGACCCCC GACTCCAAGC GCGAAAACCC CGGATGGTGA 300

GGAGCAGGCA A ATG TGC AAT ACC AAC ATG TCT GTA CCT ACT GAT GGT GCT 350

Met Cys Asn Thr Asn Met Ser Val Pro Thr Asp Gly Ala

1 5 10

GTA ACC ACC TCA CAG ATT CCA GCT TCG GAA CAA GAG ACC CTG GTT AGA 398

Val Thr Thr Ser Gln Ile Pro Ala Ser Glu Gln Glu Thr Leu Val Arg

15 20 25

CCA AAG CCA TTG CTT TTG AAG TTA TTA AAG TCT GTT GGT GCA CAA AAA 446

Pro Lys Pro Leu Leu Leu Lys Leu Leu Lys Ser Val Gly Ala Gln Lys

30 35 40 45

GAC ACT TAT ACT ATG AAA GAG GTT CTT TTT TAT CTT GGC CAG TAT ATT 494

Asp Thr Tyr Thr Met Lys Glu Val Leu Phe Tyr Leu Gly Gln Tyr Ile

50 55 60

ATG ACT AAA CGA TTA TAT GAT GAG AAG CAA CAA CAT ATT GTA TAT TGT 542

Met Thr Lys Arg Leu Tyr Asp Glu Lys Gln Gln His Ile Val Tyr Cys

65 70 75

TCA AAT GAT CTT CTA GGA GAT TTG TTT GGC GTG CCA AGC TTC TCT GTG 590

Ser Asn Asp Leu Leu Gly Asp Leu Phe Gly Val Pro Ser Phe Ser Val

80 85 90

09966724 100101

AAA GAG CAC AGG AAA ATA TAT ACC ATG ATC TAC AGG AAC TTG GTA GTA 638
Lys Glu His Arg Lys Ile Tyr Thr Met Ile Tyr Arg Asn Leu Val Val
95 100 105

GTC AAT CAG CAG GAA TCA TCG GAC TCA GGT ACA TCT GTG AGT GAG AAC 686
Val Asn Gln Gln Glu Ser Ser Asp Ser Gly Thr Ser Val Ser Glu Asn
110 115 120 125

AGG TGT CAC CTT GAA GGT GGG AGT GAT CAA AAG GAC CTT GTA CAA GAG 734
Arg Cys His Leu Glu Gly Gly Ser Asp Gln Lys Asp Leu Val Gln Glu
130 135 140

CTT CAG GAA GAG AAA CCT TCA TCT TCA CAT TTG GTT TCT AGA CCA TCT 782
Leu Gln Glu Glu Lys Pro Ser Ser Ser His Leu Val Ser Arg Pro Ser
145 150 155

ACC TCA TCT AGA AGG AGA GCA ATT AGT GAG ACA GAA GAA AAT TCA GAT 830
Thr Ser Ser Arg Arg Arg Ala Ile Ser Glu Thr Glu Glu Asn Ser Asp
160 165 170

GAA TTA TCT GGT GAA CGA CAA AGA AAA CGC CAC AAA TCT GAT AGT ATT 878
Glu Leu Ser Gly Glu Arg Gln Arg Lys Arg His Lys Ser Asp Ser Ile
175 180 185

TCC CTT TCC TTT GAT GAA AGC CTG GCT CTG TGT GTA ATA AGG GAG ATA 926
Ser Leu Ser Phe Asp Glu Ser Leu Ala Leu Cys Val Ile Arg Glu Ile
190 195 200 205

09966724-100101

TGT TGT GAA AGA AGC AGT AGC AGT GAA TCT ACA GGG ACG CCA TCG AAT	974
Cys Cys Glu Arg Ser Ser Ser Ser Glu Ser Thr Gly Thr Pro Ser Asn	
210 215 220	
CCG GAT CTT GAT GCT GGT GTA AGT GAA CAT TCA GGT GAT TGG TTG GAT	1022
Pro Asp Leu Asp Ala Gly Val Ser Glu His Ser Gly Asp Trp Leu Asp	
225 230 235	
CAG GAT TCA GTT TCA GAT CAG TTT AGT GTA GAA TTT GAA GTT GAA TCT	1070
Gln Asp Ser Val Ser Asp Gln Phe Ser Val Glu Phe Glu Val Glu Ser	
240 245 250	
CTC GAC TCA GAA GAT TAT AGC CTT AGT GAA GAA GGA CAA GAA CTC TCA	1118
Leu Asp Ser Glu Asp Tyr Ser Leu Ser Glu Glu Gly Gln Glu Leu Ser	
255 260 265	
GAT GAA GAT GAT GAG GTA TAT CAA GTT ACT GTG TAT CAG GCA GGG GAG	1166
Asp Glu Asp Asp Glu Val Tyr Gln Val Thr Val Tyr Gln Ala Gly Glu	
270 275 280 285	
AGT GAT ACA GAT TCA TTT GAA GAA GAT CCT GAA ATT TCC TTA GCT GAC	1214
Ser Asp Thr Asp Ser Phe Glu Glu Asp Pro Glu Ile Ser Leu Ala Asp	
290 295 300	
TAT TGG AAA TGC ACT TCA TGC AAT GAA ATG AAT CCC CCC CTT CCA TCA	1262
Tyr Trp Lys Cys Thr Ser Cys Asn Glu Met Asn Pro Pro Leu Pro Ser	
305 310 315	

0966724-100401

CAT TGC AAC AGA TGT TGG GCC CTT CGT GAG AAT TGG CTT CCT GAA GAT 1310
His Cys Asn Arg Cys Trp Ala Leu Arg Glu Asn Trp Leu Pro Glu Asp
320 325 330

AAA GGG AAA GAT AAA GGG GAA ATC TCT GAG AAA GCC AAA CTG GAA AAC 1358
Lys Gly Lys Asp Lys Gly Glu Ile Ser Glu Lys Ala Lys Leu Glu Asn
335 340 345

TCA ACA CAA GCT GAA GAG GGC TTT GAT GTT CCT GAT TGT AAA AAA ACT 1406
Ser Thr Gln Ala Glu Glu Gly Phe Asp Val Pro Asp Cys Lys Lys Thr
350 355 360 365

ATA GTG AAT GAT TCC AGA GAG TCA TGT GTT GAG GAA AAT GAT GAT AAA 1454
Ile Val Asn Asp Ser Arg Glu Ser Cys Val Glu Glu Asn Asp Asp Lys
370 375 380

ATT ACA CAA GCT TCA CAA TCA CAA GAA AGT GAA GAC TAT TCT CAG CCA 1502
Ile Thr Gln Ala Ser Gln Ser Gln Glu Ser Glu Asp Tyr Ser Gln Pro
385 390 395

TCA ACT TCT AGT AGC ATT ATT TAT AGC AGC CAA GAA GAT GTG AAA GAG 1550
Ser Thr Ser Ser Ser Ile Ile Tyr Ser Ser Gln Glu Asp Val Lys Glu
400 405 410

TTT GAA AGG GAA GAA ACC CAA GAC AAA GAA GAG AGT GTG GAA TCT AGT 1598
Phe Glu Arg Glu Glu Thr Gln Asp Lys Glu Glu Ser Val Glu Ser Ser
415 420 425

090624 100001

TTG CCC CTT AAT GCC ATT GAA CCT TGT GTG ATT TGT CAA GGT CGA CCT 1646
Leu Pro Leu Asn Ala Ile Glu Pro Cys Val Ile Cys Gln Gly Arg Pro
430 435 440 445

AAA AAT GGT TGC ATT GTC CAT GGC AAA ACA GGA CAT CTT ATG GCC TGC 1694
Lys Asn Gly Cys Ile Val His Gly Lys Thr Gly His Leu Met Ala Cys
450 455 460

TTT ACA TGT GCA AAG AAG CTA AAG AAA AGG AAT AAG CCC TGC CCA GTA 1742
Phe Thr Cys Ala Lys Lys Leu Lys Lys Arg Asn Lys Pro Cys Pro Val
465 470 475

TGT AGA CAA CCA ATT CAA ATG ATT GTG CTA ACT TAT TTC CCC 1784
Cys Arg Gln Pro Ile Gln Met Ile Val Leu Thr Tyr Phe Pro
480 485 490

TAGTTGACCT GTCTATAAGA GAATTATATA TTTCTAACTA TATAACCCTA GGAATTTAGA 1844

CAACCTGAAA TTTATTCACA TATATCAAAG TGAGAAAATG CCTCAATTCA CATAGATTTG 1904

TTCTCTTTAG TATAATTGAC CTACTTTGGT AGTGGAATAG TGAATACTTA CTATAATTTG 1964

ACTTGAATAT GTAGCTCATC CTTTACACCA ACTCCTAATT TTAAATAATT TCTACTCTGT 2024

CTTAAATGAG AAGTACTTGG TTTTTTTTTT CTAAATATG TATATGACAT TTAAATGTAA 2084

CTTATTATTT TTTTGTAGAC CGAGTCTTGC TCTGTTACCC AGGCTGGAGT GCAGTGGGTG 2144

ATCTTGGCTC ACTGCAAGCT CTGCCCTCCC CGGGTTCGCA CCATTCTCCT GCCTCAGCCT 2204

0966724.100101

CCCAATTAGC TTGGCCTACA GTCATCTGCC ACCACACCTG GCTAATTTT TGTACTTTTA 2264
GTAGAGACAG GGTTTCACCG TGTTAGCCAG GATGGTCTCG ATCTCCTGAC CTCGTGATCC 2324
GCCCACCTCG GCCTCCCAA GTGCTGGGAT TACAGGCATG AGCCACCG 2372

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 491 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met Cys Asn Thr Asn Met Ser Val Pro Thr Asp Gly Ala Val Thr Thr
1 5 10 15

Ser Gln Ile Pro Ala Ser Glu Gln Glu Thr Leu Val Arg Pro Lys Pro
20 25 30

Leu Leu Leu Lys Leu Leu Lys Ser Val Gly Ala Gln Lys Asp Thr Tyr
35 40 45

Thr Met Lys Glu Val Leu Phe Tyr Leu Gly Gln Tyr Ile Met Thr Lys
50 55 60

0906724-100104

Arg Leu Tyr Asp Glu Lys Gln Gln His Ile Val Tyr Cys Ser Asn Asp
65 70 75 80

Leu Leu Gly Asp Leu Phe Gly Val Pro Ser Phe Ser Val Lys Glu His
85 90 95

Arg Lys Ile Tyr Thr Met Ile Tyr Arg Asn Leu Val Val Val Asn Gln
100 105 110

Gln Glu Ser Ser Asp Ser Gly Thr Ser Val Ser Glu Asn Arg Cys His
115 120 125

Leu Glu Gly Gly Ser Asp Gln Lys Asp Leu Val Gln Glu Leu Gln Glu
130 135 140

Glu Lys Pro Ser Ser Ser His Leu Val Ser Arg Pro Ser Thr Ser Ser
145 150 155 160

Arg Arg Arg Ala Ile Ser Glu Thr Glu Glu Asn Ser Asp Glu Leu Ser
165 170 175

Gly Glu Arg Gln Arg Lys Arg His Lys Ser Asp Ser Ile Ser Leu Ser
180 185 190

Phe Asp Glu Ser Leu Ala Leu Cys Val Ile Arg Glu Ile Cys Cys Glu
195 200 205

Arg Ser Ser Ser Ser Glu Ser Thr Gly Thr Pro Ser Asn Pro Asp Leu
210 215 220

095624 100101
100000 122500

Asp Ala Gly Val Ser Glu His Ser Gly Asp Trp Leu Asp Gln Asp Ser
225 230 235 240

Val Ser Asp Gln Phe Ser Val Glu Phe Glu Val Glu Ser Leu Asp Ser
245 250 255

Glu Asp Tyr Ser Leu Ser Glu Glu Gly Gln Glu Leu Ser Asp Glu Asp
260 265 270

Asp Glu Val Tyr Gln Val Thr Val Tyr Gln Ala Gly Glu Ser Asp Thr
275 280 285

Asp Ser Phe Glu Glu Asp Pro Glu Ile Ser Leu Ala Asp Tyr Trp Lys
290 295 300

Cys Thr Ser Cys Asn Glu Met Asn Pro Pro Leu Pro Ser His Cys Asn
305 310 315 320

Arg Cys Trp Ala Leu Arg Glu Asn Trp Leu Pro Glu Asp Lys Gly Lys
325 330 335

Asp Lys Gly Glu Ile Ser Glu Lys Ala Lys Leu Glu Asn Ser Thr Gln
340 345 350

Ala Glu Glu Gly Phe Asp Val Pro Asp Cys Lys Lys Thr Ile Val Asn
355 360 365

Asp Ser Arg Glu Ser Cys Val Glu Glu Asn Asp Asp Lys Ile Thr Gln
370 375 380

09966724-100401

Ala Ser Gln Ser Gln Glu Ser Glu Asp Tyr Ser Gln Pro Ser Thr Ser
385 390 395 400

Ser Ser Ile Ile Tyr Ser Ser Gln Glu Asp Val Lys Glu Phe Glu Arg
405 410 415

Glu Glu Thr Gln Asp Lys Glu Glu Ser Val Glu Ser Ser Leu Pro Leu
420 425 430

Asn Ala Ile Glu Pro Cys Val Ile Cys Gln Gly Arg Pro Lys Asn Gly
435 440 445

Cys Ile Val His Gly Lys Thr Gly His Leu Met Ala Cys Phe Thr Cys
450 455 460

Ala Lys Lys Leu Lys Lys Arg Asn Lys Pro Cys Pro Val Cys Arg Gln
465 470 475 480

Pro Ile Gln Met Ile Val Leu Thr Tyr Phe Pro
485 490

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1710 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

0956741001
"TOPOT" 4229550

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Mus musculus

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 202..1668

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GAGGAGCCGC CGCCTTCTCG TCGCTCGAGC TCTGGACGAC CATGGTCGCT CAGGCCCCGT 60

CCGCGGGGCC TCCGCGCTCC CCGTGAAGGG TCGGAAGATG CGCGGGAAGT AGCAGCCGTC 120

TGCTGGGCGA GCGGGAGACC GACCGGACAC CCCTGGGGGA CCCTCTCGGA TCACCGCGCT 180

TCTCCTGCGG CCTCCAGGCC A ATG TGC AAT ACC AAC ATG TCT GTG TCT ACC 231

Met Cys Asn Thr Asn Met Ser Val Ser Thr

1 5 10

GAG GGT GCT GCA AGC ACC TCA CAG ATT CCA GCT TCG GAA CAA GAG ACT 279

Glu Gly Ala Ala Ser Thr Ser Gln Ile Pro Ala Ser Glu Gln Glu Thr

15 20 25

0596674100101

CTG GTT AGA CCA AAA CCA TTG CTT TTG AAG TTG TTA AAG TCC GTT GGA 327
 Leu Val Arg Pro Lys Pro Leu Leu Leu Lys Leu Leu Lys Ser Val Gly
 30 35 40

GCG CAA AAC GAC ACT TAC ACT ATG AAA GAG ATT ATA TTT TAT ATT GGC 375
 Ala Gln Asn Asp Thr Tyr Thr Met Lys Glu Ile Ile Phe Tyr Ile Gly
 45 50 55

CAG TAT ATT ATG ACT AAG AGG TTA TAT GAC GAG AAG CAG CAG CAC ATT 423
 Gln Tyr Ile Met Thr Lys Arg Leu Tyr Asp Glu Lys Gln Gln His Ile
 60 65 70

GTG TAT TGT TCA AAT GAT CTC CTA GGA GAT GTG TTT GGA GTC CCG AGT 471
 Val Tyr Cys Ser Asn Asp Leu Leu Gly Asp Val Phe Gly Val Pro Ser
 75 80 85 90

TTC TCT GTG AAG GAG CAC AGG AAA ATA TAT GCA ATG ATC TAC AGA AAT 519
 Phe Ser Val Lys Glu His Arg Lys Ile Tyr Ala Met Ile Tyr Arg Asn
 95 100 105

TTA GTG GCT GTA AGT CAG CAA GAC TCT GGC ACA TCG CTG AGT GAG AGC 567
 Leu Val Ala Val Ser Gln Gln Asp Ser Gly Thr Ser Leu Ser Glu Ser
 110 115 120

AGA CGT CAG CCT GAA GGT GGG AGT GAT CTG AAG GAT CCT TTG CAA GCG 615
 Arg Arg Gln Pro Glu Gly Gly Ser Asp Leu Lys Asp Pro Leu Gln Ala
 125 130 135

09565724 100103

CCA CCA GAA GAG AAA CCT TCA TCT TCT GAT TTA ATT TCT AGA CTG TCT 663
Pro Pro Glu Glu Lys Pro Ser Ser Ser Asp Leu Ile Ser Arg Leu Ser
140 145 150

ACC TCA TCT AGA AGG AGA TCC ATT AGT GAG ACA GAA GAG AAC ACA GAT 711
Thr Ser Ser Arg Arg Arg Ser Ile Ser Glu Thr Glu Glu Asn Thr Asp
155 160 165 170

GAG CTA CCT GGG GAG CGG CAC CGG AAG CGC CGC AGG TCC CTG TCC TTT 759
Glu Leu Pro Gly Glu Arg His Arg Lys Arg Arg Arg Ser Leu Ser Phe
175 180 185

GAT CCG AGC CTG GGT CTG TGT GAG CTG AGG GAG ATG TGC AGC GGC GGC 807
Asp Pro Ser Leu Gly Leu Cys Glu Leu Arg Glu Met Cys Ser Gly Gly
190 195 200

ACG AGC AGC AGT AGC AGC AGC AGC AGC AGC GAG TCC ACA GAG ACG CCC TCG 855
Thr Ser Ser Ser Ser Ser Ser Ser Ser Ser Glu Ser Thr Glu Thr Pro Ser
205 210 215

CAT CAG GAT CTT GAC GAT GGC GTA AGT GAG CAT TCT GGT GAT TGC CTG 903
His Gln Asp Leu Asp Asp Gly Val Ser Glu His Ser Gly Asp Cys Leu
220 225 230

GAT CAG GAT TCA GTT TCT GAT CAG TTT AGC GTG GAA TTT GAA GTT GAG 951
Asp Gln Asp Ser Val Ser Asp Gln Phe Ser Val Glu Phe Glu Val Glu
235 240 245 250

05966724-100101

TCT CTG GAC TCG GAA GAT TAC AGC CTG AGT GAC GAA GGG CAC GAG CTC 999
 Ser Leu Asp Ser Glu Asp Tyr Ser Leu Ser Asp Glu Gly His Glu Leu
 255 260 265

TCA GAT GAG GAT GAT GAG GTC TAT CGG GTC ACA GTC TAT CAG ACA GGA 1047
 Ser Asp Glu Asp Asp Glu Val Tyr Arg Val Thr Val Tyr Gln Thr Gly
 270 275 280

GAA AGC GAT ACA GAC TCT TTT GAA GGA GAT CCT GAG ATT TCC TTA GCT 1095
 Glu Ser Asp Thr Asp Ser Phe Glu Gly Asp Pro Glu Ile Ser Leu Ala
 285 290 295

GAC TAT TGG AAG TGT ACC TCA TGC AAT GAA ATG AAT CCT CCC CTT CCA 1143
 Asp Tyr Trp Lys Cys Thr Ser Cys Asn Glu Met Asn Pro Pro Leu Pro
 300 305 310

TCA CAC TGC AAA AGA TGC TGG ACC CTT CGT GAG AAC TGG CTT CCA GAC 1191
 Ser His Cys Lys Arg Cys Trp Thr Leu Arg Glu Asn Trp Leu Pro Asp
 315 320 325 330

GAT AAG GGG AAA GAT AAA GTG GAA ATC TCT GAA AAA GCC AAA CTG GAA 1239
 Asp Lys Gly Lys Asp Lys Val Glu Ile Ser Glu Lys Ala Lys Leu Glu
 335 340 345

AAC TCA GCT CAG GCA GAA GAA GGC TTG GAT GTG CCT GAT GGC AAA AAG 1287
 Asn Ser Ala Gln Ala Glu Glu Gly Leu Asp Val Pro Asp Gly Lys Lys
 350 355 360

09956724 "100104"

CTG ACA GAG AAT GAT GCT AAA GAG CCA TGT GCT GAG GAG GAC AGC GAG 1335
 Leu Thr Glu Asn Asp Ala Lys Glu Pro Cys Ala Glu Glu Asp Ser Glu
 365 370 375

GAG AAG GCC GAA CAG ACG CCC CTG TCC CAG GAG AGT GAC GAC TAT TCC 1383
 Glu Lys Ala Glu Gln Thr Pro Leu Ser Gln Glu Ser Asp Asp Tyr Ser
 380 385 390

CAA CCA TCG ACT TCC AGC AGC ATT GTT TAT AGC AGC CAA GAA AGC GTG 1431
 Gln Pro Ser Thr Ser Ser Ser Ile Val Tyr Ser Ser Gln Glu Ser Val
 395 400 405 410

AAA GAG TTG AAG GAG GAA ACG CAG CAC AAA GAC GAG AGT GTG GAA TCT 1479
 Lys Glu Leu Lys Glu Glu Thr Gln His Lys Asp Glu Ser Val Glu Ser
 415 420 425

AGC TTC TCC CTG AAT GCC ATC GAA CCA TGT GTG ATC TGC CAG GGG CGG 1527
 Ser Phe Ser Leu Asn Ala Ile Glu Pro Cys Val Ile Cys Gln Gly Arg
 430 435 440

CCT AAA AAT GGC TGC ATT GTT CAC GGC AAG ACT GGA CAC CTC ATG TCA 1575
 Pro Lys Asn Gly Cys Ile Val His Gly Lys Thr Gly His Leu Met Ser
 445 450 455

TGT TTC ACG TGT GCA AAG AAG CTA AAA AAA AGA AAC AAG CCC TGC CCA 1623
 Cys Phe Thr Cys Ala Lys Lys Leu Lys Lys Arg Asn Lys Pro Cys Pro
 460 465 470

096674 100101

GTG TGC AGA CAG CCA ATC CAA ATG ATT GTG CTA AGT TAC TTC AAC 1668
Val Cys Arg Gln Pro Ile Gln Met Ile Val Leu Ser Tyr Phe Asn
475 480 485

TAGCTGACCT GCTCACAAAA ATAGAATTTT ATATTTCTAA CT 1710

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 489 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met Cys Asn Thr Asn Met Ser Val Ser Thr Glu Gly Ala Ala Ser Thr
1 5 10 15

Ser Gln Ile Pro Ala Ser Glu Gln Glu Thr Leu Val Arg Pro Lys Pro
20 25 30

Leu Leu Leu Lys Leu Leu Lys Ser Val Gly Ala Gln Asn Asp Thr Tyr
35 40 45

Thr Met Lys Glu Ile Ile Phe Tyr Ile Gly Gln Tyr Ile Met Thr Lys
50 55 60

0956724-100101

Leu Leu Gly Asp Val Phe Gly Val Pro Ser Phe Ser Val Lys Glu His
85 90 95

Arg Lys Ile Tyr Ala Met Ile Tyr Arg Asn Leu Val Ala Val Ser Gln
100 105 110

Gln Asp Ser Gly Thr Ser Leu Ser Glu Ser Arg Arg Gln Pro Glu Gly
115 120 125

Gly Ser Asp Leu Lys Asp Pro Leu Gln Ala Pro Pro Glu Glu Lys Pro
130 135 140

Ser Ser Ser Asp Leu Ile Ser Arg Leu Ser Thr Ser Ser Arg Arg Arg
145 150 155 160

Ser Ile Ser Glu Thr Glu Glu Asn Thr Asp Glu Leu Pro Gly Glu Arg
165 170 175

His Arg Lys Arg Arg Arg Ser Leu Ser Phe Asp Pro Ser Leu Gly Leu
180 185 190

Cys Glu Leu Arg Glu Met Cys Ser Gly Gly Thr Ser Ser Ser Ser Ser
195 200 205

Ser Ser Ser Glu Ser Thr Glu Thr Pro Ser His Gln Asp Leu Asp Asp
210 215 220

Lys Glu Pro Cys Ala Glu Glu Asp Ser Glu Glu Lys Ala Glu Gln Thr
370 375 380

Ser Ile Val Tyr Ser Ser Gln Glu Ser Val Lys Glu Leu Lys Glu Glu
405 410 415

Ile Glu Pro Cys Val Ile Cys Gln Gly Arg Pro Lys Asn Gly Cys Ile
435 440 445

Lys Leu Lys Lys Arg Asn Lys Pro Cys Pro Val Cys Arg Gln Pro Ile
465 470 475 480

Gln Met Ile Val Leu Ser Tyr Phe Asn